

COOLEY LLP
BOBBY GHAJAR (198719)
(bghajar@cooley.com)
COLETTE GHAZARIAN (322235)
(cghazarian@cooley.com)
1333 2nd Street, Suite 400
Santa Monica, California 90401
Telephone: (310) 883-6400

MARK WEINSTEIN (193043)
(mweinstein@cooley.com)
KATHLEEN HARTNETT (314267)
(khartnett@cooley.com)
JUDD LAUTER (290945)
(jlauter@cooley.com)
Elizabeth L. Stameshkin (260865)
(lstameshkin@cooley.com)
3175 Hanover Street
Palo Alto, CA 94304-1130
Telephone: (650) 843-5000

LEX LUMINA PLLC
MARK A. LEMLEY (155830)
(mlemley@lex-lumina.com)
745 Fifth Avenue, Suite 500
New York, NY 10151
Telephone: (646) 898-2055

CLEARY GOTTlieb STEEN & HAMILTON LLP
ANGELA L. DUNNING (212047)
(adunning@cgsh.com)
1841 Page Mill Road, Suite 250
Palo Alto, CA 94304
Telephone: (650) 815-4131

Counsel for Defendant Meta Platforms, Inc.

UNITED STATES DISTRICT COURT
NORTHERN DISTRICT OF CALIFORNIA
SAN FRANCISCO DIVISION

RICHARD KADREY, *et al.*,

Individual and Representative Plaintiffs,

v.

META PLATFORMS, INC., a Delaware
corporation;

Defendant.

Case No. 3:23-cv-03417-VC-TSH

DECLARATION OF KATHLEEN HARTNETT

Trial Date: None
Date Action Filed: July 7, 2023

1 I, Kathleen Hartnett, declare:

2 1. I am a Partner at the law firm of Cooley LLP and counsel to Meta Platforms, Inc. in
3 the above-referenced matter. I have personal knowledge of the facts contained in this Declaration
4 and, if called as a witness, could competently testify to them under oath.

5 2. Attached as Exhibit 1 is a true and correct copy of a letter dated September 27, 2024,
6 from myself to Plaintiffs' counsel of record, regarding Plaintiffs' request to modify the Court-
7 ordered case schedule, and Exhibits A and B to this letter.

8 3. Attached as Exhibit 2 is a true and correct copy of an email dated September 27,
9 2024, from Maxwell Pritt to Cole Poppell and myself.

10 4. I understand that Meta has produced source code, including change history and
11 comments, that was used to create the Llama models. This source code was collected earlier this
12 year, after the release of Llama 1 and 2, and after much of the pre-training for Llama 3 had already
13 occurred. Meta made this code available for review on May 10, 2024. Plaintiffs first sought to
14 review this code in August of 2024. I also understand that, as a courtesy, Meta is in the process of
15 supplementing this source code production to account for the time between collection and the
16 present. Plaintiffs raise a number of vague allegations about this code, including that it is a "hodge
17 podge of code and scripts," when Meta in fact produced specific source code repositories retaining
18 the file structure and organization of the repositories used at Meta. Plaintiffs have not met and
19 conferred regarding any concern regarding a "hodge podge" of "code and scripts" or code that
20 relates to what they term "how LLMs handle or filter copyrighted material."

21 5. I understand that Meta has provided a list of and produced all pretraining and post
22 training data (i.e., data for fine-tuning and alignment) for Llama 1 and Llama 2, much of which was
23 also used for Llama 3. Meta has also supplemented its production with Llama 3 training data
24 relating to books (i.e., the "Libgen" dataset) pursuant to Rule 26. Collection of this training data
25 involves numerous technical hurdles, and the training data used to train these models comprises
26 terabytes of irrelevant data (e.g., data not relating to books). The parties held a meet and confer on
27 August 29, 2024, where they resolved that Meta need not produce additional Llama 3 data beyond
28 the "Libgen" dataset.

1 6. I understand that there was a typographical error in the September 10, 2024, cover
2 letter accompanying Meta's production of META_KADREY_DATA_005, which completed
3 Meta's production of the Libgen dataset. Plaintiffs referred to META_KADREY_DATA_005 in
4 their Rule 30(b)(6) notice served just six days after Plaintiffs received the hard drive containing
5 META_KADREY_DATA_005.

6 7. Plaintiffs state in their motion that Meta has not produced "documents that Meta
7 obtained from the totality of shadow libraries and online sources." Dkt. 193 at 4. I understand that
8 Meta performed a search for communications with certain online sources pursuant to Plaintiffs'
9 discovery requests, and, as noted above, has produced relevant data used to train the Llama 1, 2,
10 and 3 models.

11 8. I understand that Meta has also provided an interrogatory response and non-
12 privileged documents that relate to its mitigation work to prevent training data memorization.

13 9. I understand that Meta has produced documents relating to the "Libgen" dataset and,
14 as noted above, the Libgen data used to train the Llama 3 model.

15 10. On September 30, 2024, Meta responded to Plaintiffs' requests for production
16 relating to licensing, including Request for Production No. 77.

17 11. On September 20, 2024, Plaintiffs informed Meta that they intended to take the
18 deposition of Dr. Yann LeCun in place of Ms. Amanda Kallet (after previously confirming Ms.
19 Kallet's deposition for September 27, 2024).

20 12. Meta and Plaintiffs' counsel have met and conferred regarding concerns around both
21 parties' privilege logs, and Meta plans to provide an updated privilege log to Plaintiffs on Monday,
22 October 7, 2024.

23 13. At the time that Plaintiffs filed their motion, Meta had not yet responded to
24 Plaintiffs' Interrogatories regarding fair use and advice of counsel, as those interrogatory responses
25 were due that day, September 30, 2024. Meta timely served its responses on September 30, 2024.

26 14. I understand that Meta has produced documents responsive to requests for
27 production served by Plaintiffs for financial information relating to the Llama models. Meta has
28 also identified a witness on this topic, Amrish Acharya, Meta's Senior Director of Finance.

